

The Carbon Copy onto Dirty Paper Channel with Statistically Equivalent States

Stefano Rini * and Shlomo Shamai (Shitz) †

* National Chiao-Tung University, Hsinchu, Taiwan

E-mail: stefano@nctu.edu.tw

† Technion-Israel Institute of Technology, Haifa, Israel

E-mail: sshlomo@ee.technion.ac.il

Abstract—Costa’s “writing on dirty paper” capacity result establishes that full state pre-cancellation can be attained in Gel’fand-Pinsker channel with additive state and additive Gaussian noise. The “carbon copy onto dirty paper” channel is the extension of Costa’s model to the compound setting: M receivers each observe the sum of the channel input, Gaussian noise and one of M Gaussian state sequences and attempt to decode the same common message. The state sequences are all non-causally known at the transmitter which attempts to simultaneously pre-code its transmission against the channel state affecting each output. In this correspondence we derive the capacity to within 2.25 bits-per-channel-use of the carbon copying onto dirty paper channel in which the state sequences are statistically equivalent, having the same variance and the same pairwise correlation. For this channel capacity is approached by letting the channel input be the superposition of two codewords: a base codeword, simultaneously decoded at each user, and a top codeword which is pre-coded against the state realization at each user for a portion $1/M$ of the time. The outer bound relies on a recursive bounding in which incremental side information is provided at each receiver. This result represents a significant first step toward determining the capacity of the most general “carbon copy onto dirty paper” channel in which state sequences appearing in the different channel outputs have any jointly Gaussian distribution.

Index Terms—Gel’fand-Pinsker Problem; Compound State-Dependent Channel; Carbon Copying onto Dirty Paper;

In the Gel’fand-Pinsker (GP) channel [1] the output of a point-to-point channel is obtained as a random function of the channel input and a state sequence which is provided non-causally to the encoder but is unknown at the decoder. Costa’s “Writing on Dirty Paper” (WDP) channel [2] is the Gaussian version of the GP channel in which the channel output is obtained as a linear combination of the input, the state sequence and iid, Gaussian-distributed, noise. Perhaps surprisingly, Costa showed that the capacity of the WDP channel is the same as the capacity of the point-to-point channel in which the state is not present in the channel output. In other words, the transmitter can fully pre-code its transmissions against the channel state and thus the presence of the channel state does not affect capacity. The “Carbon Copying onto Dirty Paper” (CCDP) channel [3] is the extension of the GP channel to the compound scenario: in this model the transmitter wishes to communicate the same message to M receivers which observe as channel output the summation of the channel input, iid Gaussian noise and one of M state sequences. These state sequences are all provided non-causally to the transmitter but

are unknown at the receivers.

In this correspondence we derive the capacity of the CCDP channel with any number of users for the case in which the states are statically equivalent, being Gaussian-distributed with the same variance and the same pairwise correlation. We first show the approximate capacity for the case of $M = 2$ and independent channel states, then generalize this result for the case of any M and independent channel states and, lastly, show the approximate capacity for any M and any correlation.

The CCDP is a special case of the compound GP channel for which, unfortunately, not many results are available in the literature. An achievable region for the two-user compound GP channel is presented in [4] where it is shown that using a common message potentially improves over extensions of the capacity achieving strategy for the GP channel in which the channel input is pre-coded against both channel states. The CCDP was first proposed in [3] where the authors consider both the binary and the Gaussian versions of the M -user compound GP channel and derive the first inner and outer bounds for these models. We have previously considered the case of two users in [5] and derived the approximate capacity for a certain set of correlations among circularly-symmetric Gaussian state sequences. A model related to the CCDP channel is the state-dependent broadcast channel with a common message. This model is obtained from the CCDP channel by introducing two private messages to be communicated between the transmitter and each of the users. A first achievable region for this channel is obtained in [6] combining coding strategies for the GP channel and the broadcast channel [7]. Steinberg in [8] studies the channel in which one of the users is provided with the state sequence while the other user observes a degraded channel output: capacity for this channel is obtained using bounding techniques inspired by the proof of the degraded broadcast channel capacity.

The remainder of the paper is organized as follows: in Sec. I we introduce the channel model, in Sec. II we present the relevant results available in the literature. In Sec. III we derive the approximate capacity for the case $M = 2$ and independent channel states while in Sec. IV we present the approximate capacity for the case of any M and independent channel states. In Sec. V we present the approximate capacity of for the case of any pairwise correlation. Sec. VI concludes the paper.

Only sketches of the proofs are provided in the main text; the full proofs can be found in appendix.

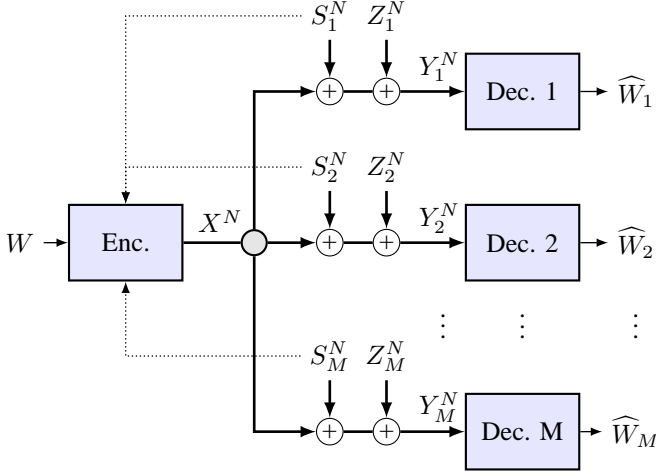


Fig. 1: The “Carbon Copying on Dirty Paper” (CCDP).

I. CHANNEL MODEL

The M -user “Carbon Copying on Dirty Paper” (CCDP) channel, also depicted in Fig. 1, is the compound GP channel in which the channel outputs are obtained as

$$Y_m^N = X^N + cS_m^N + Z_m^N, \quad m \in [1 \dots M], \quad (1)$$

where Z_m^N , $\forall m$ is an iid Gaussian sequence with zero mean and unitary variance and $\{S_m^N, m \in [1 \dots M]\}$ is an iid jointly Gaussian sequence with zero mean and covariance matrix Σ_S with

$$1 = \text{Var}[S_1] \leq \text{Var}[S_2] \dots \leq \text{Var}[S_M], \quad (2)$$

where (2) is assumed without loss of generality. The transmitter has anti-causal knowledge of $\{S_m^N, m \in [1 \dots M]\}$ and is subject to the average power constraint $\sum_{n=1}^N \mathbb{E}[|X_n|^2] \leq NP$.

In the following we focus on the CCDP in which each state has unitary variance and each two states have the same correlation. We term this model as “Carbon Copying on Dirty Paper with Equivalent States” (CCDP-ES), since all the channel states are statistically equivalent. The range of feasible values for the correlation ρ is shown by the next lemma.

Lemma I.1. *Let the matrix Σ_S be equal to*

$$\Sigma_S = (1 - \rho)\mathbf{I}_{M,M} + \rho\mathbf{1}_{M,M} = \begin{bmatrix} 1 & \rho & \dots \\ \rho & 1 & \ddots \\ \vdots & \ddots & \ddots \end{bmatrix}, \quad (3)$$

where $\mathbf{I}_{M,M}$ is the identity matrix of size M and $\mathbf{1}_{M,M}$ is the matrix of all ones of size $M \times M$, then Σ_S is positive defined for

$$-1/(M-1) \leq \rho \leq 1. \quad (4)$$

Proof: See App. A. ■

Lemma I.2. *The capacity of the CCDP channel is decreasing in c .*

Proof: See App. B. ■

This result is rather intuitive since capacity can only increase if we reduce the variance of the state.

II. RELATED RESULTS

• **Carbon Copy onto Dirty Paper (CCDP) channel.** The channel model in (1) was originally introduced in [3], in which the authors derive a number of inner and outer bounds to capacity.

Theorem II.1. *Inner and outer bounds for the 2-CCDP channel with independent states [3, Th. 3, Th. 4]. Consider the CCDP channel in (1) for $M = 2$ and $\Sigma_S = \mathbf{I}_{2,2}$, then capacity is upper bounded as*

$$\mathcal{C} \leq R^{\text{OUT}} = \begin{cases} \frac{1}{4} \log \left(\frac{1+P}{c^2/4+1} \right) \\ \quad + \frac{1}{4} \left(\frac{1+P+c^2+2c\sqrt{P}}{c^2/4+1} \right) & c^2 < 4 \\ \frac{1}{4} \log(1+P) - \frac{1}{4} \log(c^2) \\ \quad + \frac{1}{4} \log(1+P+c^2+2c\sqrt{P}) & c^2 \geq 4 \end{cases} \quad (5)$$

and lower bounded as

$$\mathcal{C} \geq R^{\text{IN}} = \begin{cases} \frac{1}{2} \log \left(1 + \frac{P}{c^2/2+1} \right) & c^2/2 \leq 1 \\ \frac{1}{2} \log \left(\frac{P+c^2/2+1}{c^2} \right) \\ \quad + \frac{1}{4} \log \left(\frac{c^2}{2} \right) & 1 \leq c^2/2 < P+1 \\ \frac{1}{4} \log(P+1) & c^2/2 \geq P+1 \end{cases} \quad (6)$$

A powerful bounding techniques is introduced in [3] to derive the outer bound in (5) while the inner bound in (6) is obtained by having the transmitter pre-code against two linear combinations of the state sequences.

The outer bounding technique for the case of $M = 2$ is also extended to the case of a general M .

Theorem II.2. *Outer bounds for the M-CCDP channel with independent states [3, Eq. (31)]. Consider the CCDP channel in (1) for and $\Sigma_S = \mathbf{I}_{M,M}$, then capacity can be upper bounded as*

$$\mathcal{C} \leq R^{\text{OUT}} = \frac{1}{2} \log(P + c^2 + 2c\sqrt{P}) - \frac{M-1}{2M} \log c^2 \\ - \frac{1}{2M} \log M - \left[\frac{1}{2M} \log \left(\frac{c^2}{M(P+1)} \right) \right]^+. \quad (7)$$

Inner and outer bounds for the case $M = 2$ are close for small values of P but otherwise no capacity characterization is possible using the bounds in Th. II.1. By generalizing the inner bound in (6) to any M , we can again show that inner and outer bound are close only for small values of P .

• **Compound GP.** The compound GP is a more general channel model than the CCDP: in [4] an attainable rate region for this model is obtained as:

$$R^{\text{IN}} \leq \max_{P_{X,V,U_1,U_2}} \min \left\{ I_1, I_2, \frac{1}{2} (I_1 + I_2 - I(U_1; U_2 | V, S_1, S_2)) \right\}, \quad (8)$$

for $I_i = I(Y_i; U_i, V) - I(V, U_i; S_1, S_2)$, $i \in \{1, 2\}$. The variable V is a common message decoded at both receivers, while U_1 and U_2 are pre-coded against S_1 and S_2 respectively as in the GP channel.

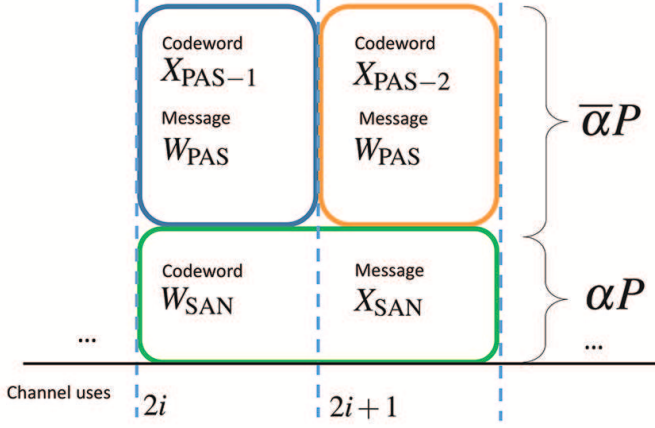


Fig. 2: A graphical representation of the capacity approaching scheme in Th. III.1.

III. THE 2-CCDP CHANNEL WITH INDEPENDENT, EQUAL-VARIANCE STATES

We begin by deriving the approximate capacity for 2-CCDP-ES for $\rho = 0$: this allows us to illustrate the main inner and outer bounding techniques while deferring more complex derivations to the latter sections. In the derivation of the inner bound, we consider the same attainable strategy as in [9], also depicted in Fig. 2: the channel input is obtained as the superposition of three codewords: (i) a bottom common codeword, X_{SAN}^N (SAN for *State As Noise*) with power αP , carries the message W_{SAN} with rate R_{SAN} and treats the state sequences S_1^N and S_2^N as additional noise while, and (ii) two top private codewords, $X_{\text{PAS}-1}^N, X_{\text{PAS}-2}^N$ (PAS - i for *Pre-coded Against State S_i^N*), with power $\bar{\alpha}P$ for $\bar{\alpha} = 1 - \alpha$, pre-coded against S_1^N and S_2^N respectively and transmitted for half of the time each. Since the $\text{Var}[S_1] = \text{Var}[S_2]$, the codeword X_{SAN}^N can be decoded at both receivers simultaneously. On the other hand, $X_{\text{PAS}-i}^N$ is decoded only at receiver i since it is pre-coded against the state S_i^N . In order for the both decoders to decode the same amount of common information, these codewords carry the same message W_{PAS} at rate R_{PAS} . As a result of these consideration, both receivers are able to correctly decode both W_{SAN} and W_{PAS} , thus attaining the transmission rate

$$R^{\text{IN}} = \frac{1}{2} \log \left(1 + \frac{\alpha P}{c^2 + \bar{\alpha} P + 1} \right) + \frac{1}{4} \log (1 + \bar{\alpha} P). \quad (9)$$

The expression in (9) can be maximized over α , the ratio between the power of the common and the private codewords. When $P+1 \geq c^2$, the optimal value of $\bar{\alpha}$ is $(c^2 - 1)/P$, which corresponds to fixing the power of the private codewords to the same power as the state sequence. When $c^2 > P+1$, instead, all the power is allocated to the private codewords and the scheme reduces to pre-coding for receiver 1 half of the time and pre-coding for receiver 2 the remaining portion of the time.

With respect to the outer bound, we are able to improve on the result of Th. II.1 using the observation in Lem. I.2: note that the outer bound expression in (5) for $c > 4$ is not

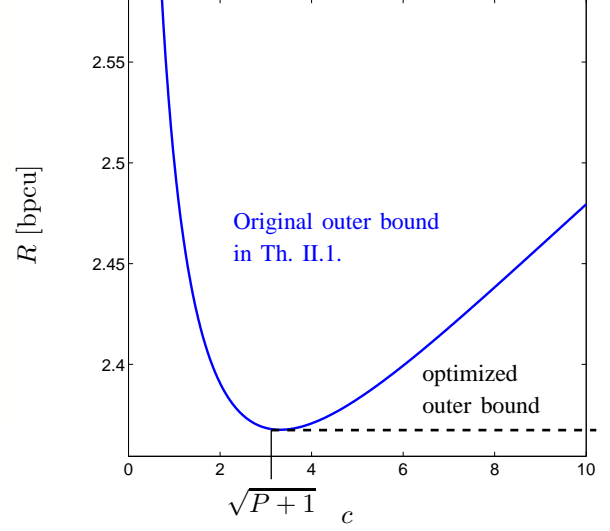


Fig. 3: The outer bound in (5) for $P = 10$ and $c \in [0, 10]$.

decreasing increasing in c , as shown Fig. 3. For this reason it is possible to improve the outer bound by considering a channel with a parameter $c' = \min\{\sqrt{P+1}, c\} \leq c$: this channel has a larger capacity than the original channel but provides a tighter outer bound. By comparing these inner and outer bound expressions, we can bound the capacity to within 1 bpcu.

Theorem III.1. Approximate capacity for the 2-CCDP with independent, equal-variance states.

Consider the 2-CCDP-IS channel in Fig. 1 for $\Sigma_S = \mathbf{I}_{2,2}$, then an outer bound to capacity is

$$\mathcal{C} \leq R^{\text{OUT}} = \begin{cases} \frac{1}{2} \log(P+1) & c^2 \leq 1 \\ \frac{1}{2} \log(P+c^2+1) & 1 < c^2 < P+1 \\ -\frac{1}{4} \log(c^2+1) + \frac{1}{2} & c^2 \geq P+1 \end{cases} \quad (10)$$

and the exact capacity \mathcal{C} is to within a gap of 1 bpcu from the outer bound in (10).

Proof: See. App. C. ■

The result in Th. III.1 is somewhat expected: when the states in the 2-CCDP channel are independent, the best strategy is to send a common codeword at a power level larger than the channel state that can be decoded at both users and a private codeword for each user, pre-coded against the state realization in the corresponding channel output. In order for the private codeword to communicate the same message at the two receiver, this codeword must be time-shared between the two receivers. The major difficulty in proving theorem is therefore in deriving an outer bound which matches this intuitively optimal solution. Before showing the approximate capacity of the CCDP-ES, we first show how to extend of the result it Thm. III.1 from the case of $M = 2$ to the case of any number of users.

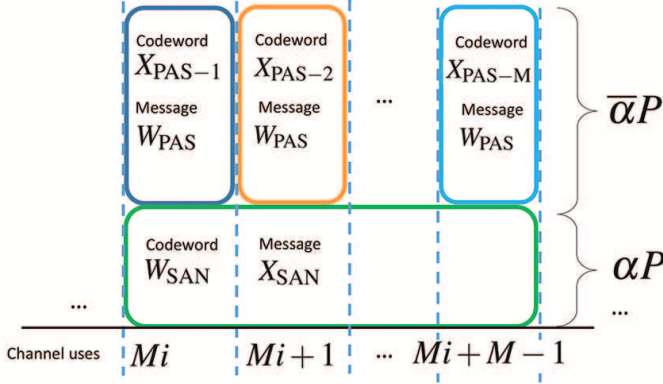


Fig. 4: A graphical representation of the capacity approaching scheme in Th. III.1.

IV. THE M-CCDP CHANNEL WITH INDEPENDENT, EQUAL-VARIANCE STATES

The approximate capacity of the M-CCDP channel with independent, equal-variance states is obtained through the appropriate extension of the inner and outer bounds in Sec. III. A generalization of the inner bound in Fig. 2 to the case of any number of users is rather straightforward: we can modify the attainable strategy in Fig. 2 as shown in Fig. 4 and employ one common codeword X_{SAN}^N at power αP and M time-shared codewords $X_{\text{PAS}-m}^N$, $m \in [1 \dots M]$ of power $\bar{\alpha}P$, each pre-coded against the state sequence S_m^N . All the codewords $X_{\text{PAS}-m}^N$ convey the same message W_{PAS} and receiver m decodes both the codeword X_{SAN}^N and $X_{\text{PAS}-m}^N$ so that, at the end of the transmission, all the decoders can correctly decode both W_{SAN} and W_{PAS} . The rate that we can attain with this strategy is

$$R^{\text{IN}} = \frac{1}{2} \log \left(1 + \frac{\alpha P}{c^2 + \bar{\alpha}P + 1} \right) + \frac{1}{2M} \log (1 + \bar{\alpha}P), \quad (11)$$

which can again be maximized over α . In this case the optimal value of $\bar{\alpha}$ is

$$\bar{\alpha}^* = \max \left\{ 0, \min \left\{ 1, \frac{c^2 + 1 - M}{P(M-1)} \right\} \right\}, \quad (12)$$

and the above scheme reduces to simple time-sharing and Costa pre-coding when $c^2 > (M-1)(P+1)$.

The generalization of the outer bound in Th. III.1 is rather more involved: this can be accomplished by establishing a recursive bounding of the mutual information terms obtained from Fano's inequality and using a very carefully-chosen genie side information for each decoder. We refer the interested reader to [10, App. D] for the complete proof. Again, the observation in Lem. I.2 is employed to tighten the outer bound expression by optimizing over the state gain c .

Theorem IV.1. Approximate capacity M-user CCDP with independent, equal-variance states.

Consider the M-CCDP-IS channel in Fig. 1 for $\Sigma_S = \mathbf{I}_{M,M}$ then an outer bound to capacity is

$$\mathcal{C} \leq R^{\text{OUT}} =$$

$$\begin{cases} \frac{1}{2} \log \left(1 + \frac{P}{1+c^2} \right) + \frac{9}{4} & M-1 \geq c^2 \\ \frac{1}{2M} \log(1+P) + \frac{M-1}{2M} \log(c^2) + \frac{3}{2} & M-1 < c^2 \leq (M-1)(P+1) \\ \frac{1}{2M} \log(1+P) + 2 & c^2 > (M-1)(P+1) \end{cases} \quad (13)$$

and the exact capacity \mathcal{C} is to within a gap of 2.25 bpcu from the outer bound in (13).

Proof: See App. D. ■

It is interesting to notice that pure time-sharing with no common codeword is approximatively optimal when $c^2 > (M-1)(P+1)$ that is when the state variance is roughly M times stronger than the transmit power. This occurs, intuitively, because the pre-log of the rate of the codeword X_{SAN}^N is $1/2$ while the pre-log of the codewords $X_{\text{PAS}-m}^N$ is $1/2M$.

V. THE CCDP-ES CHANNEL

In this section we finally derive the approximate capacity of the CCDP-ES channel: the result relies, from a high-level viewpoint, on two observations: (i) positive correlation among the states implies that there exists a common component which can be pre-coded against in the common codeword X_{SAN}^N , and (ii) negative correlation among the states does not allow any improvement in the attainable rates with respect to the case of independent channel states. To illustrate these points, note that the output of the 2-CCDP-ES can be equivalently expressed as

$$Y_1^N = X + c \left(a S_c^N + \sqrt{1-a} \tilde{S}_1^N \right) + Z_1^N \quad (14a)$$

$$Y_2^N = X + c \left(\frac{\rho}{a} S_c^N + \sqrt{1-\frac{\rho^2}{a^2}} \tilde{S}_2^N \right) + Z_2^N, \quad (14b)$$

for some $S_c, \tilde{S}_1, \tilde{S}_2 \sim \mathcal{N}(0,1)$, *iid*, and any $a \in [-1, +1]$. The choice $a = \sqrt{|\rho|}$ makes the term S_c have the same scaling in both channel outputs: for the case of positive correlation this term can be simultaneously pre-coded at both receivers as in the WDP channel. For of negative correlation, since the common term appears in with opposite sign in the two outputs, no coding advantage is possible.

Theorem V.1. Approximate capacity for the general 2-CCDP-ES.

Consider the general 2-CCDP channel with state covariance matrix Σ_S as in (3) for ρ satisfying (4), then capacity can be upper bounded as

$$\mathcal{C} \leq R^{\text{OUT}} = \begin{cases} \frac{1}{2} \log(P+1) & c^2 \bar{\rho}^+ \leq 1 \\ \frac{1}{2} \log(P + c^2 \bar{\rho}^+ + 1) & 1 < c^2 \bar{\rho}^+ < P+1 \\ -\frac{1}{4} \log(\bar{\rho}^+ c^2) + \frac{1}{2} & c^2 \bar{\rho} \geq P+1 \end{cases} \quad (15)$$

for $\bar{\rho}^+ = 1 - \max\{\rho, 0\}$ and the exact capacity is to within 2.25 bpcu from the outer bound in (16).

Proof: See App. E. ■

The outer bound in (16) for $\rho > 0$ is obtained by providing the common state S_c as a side information to the receiver: the resulting channel is then the same model as in Th. III.1

but with $c' = \bar{\rho}c$. For the case of $\rho < 0$, we rely on the fact that outer bound in Th. III.1, when adapted to the case of correlated states, is increasing in the parameter ρ and thus the case of $\rho = 0$ provides a looser outer bound than the case of $\rho < 0$. The achievability proof for the case $\rho < 0$ is the same as the achievability proof in Th. III.1, since this scheme is not affected by correlation among the states. For the case of $\rho > 0$ we adapt the scheme in Th. III.1 by having the common codeword X_{SAN}^N pre-coded against the common state sequence $c\sqrt{\bar{\rho}}S_c^N$.

The decomposition of the channel outputs in (14) in terms of a common component can be extended to the case of any users, and the distinction between positive and negative pairwise correlation becomes clearer in this context.

For the case of positive correlation, a common term with variance ρ can be extracted from all channel outputs by representing the channel states as

$$S_m = \sqrt{\rho}S_c + \sqrt{1-\rho}\tilde{S}_m \quad m \in [1 \dots M], \quad (17)$$

for $S_c, \tilde{S}_m \sim \mathcal{N}(0,1)$, *iid*. As for the proof of Th. V.1, the transmitter can simultaneously pre-code against the term $\sqrt{\rho}S_c$ at all the users as in the WDP channel.

The case of negative correlation is more intriguing, since, in this case, the channel states can be represented as

$$S_m = \sum_{j=m+1}^N \sqrt{\rho}\hat{S}_{mj} - \sum_{j=1}^{m-1} \sqrt{\rho}\hat{S}_{jm} + \sqrt{1-(N-1)\rho}\tilde{S}_m, \quad (18)$$

for $\hat{S}_{mj}, \tilde{S}_m \sim \mathcal{N}(0,1)$, $[m,j] \in [1 \dots M]^2$, $m > j$. The representation in (18) provides some intuition on the result in Lem. I.1: in order for the two states, S_j and S_k with $k > j$, to be negatively correlated, they must share a term \hat{S}_{jk} that does not appear in any other S_m . This must be the case, otherwise this term would affect the correlation among S_j, S_k and S_m . Since each S_m must be negatively correlated with other $N-1$ states, it must contain $N-1$ terms \hat{S}_{mj} or \hat{S}_{jm} , each with variance $|\rho|$. Given that the variance of S_m is equal to one, we necessarily have that $|\rho| \leq 1/(N-1)$ or $\rho > -1/(N-1)$. With the considerations in (17) and (18) we can finally state the main result of the paper.

Theorem V.2. Approximate capacity for the M-CCDP-ES. Consider the M-CCDP channel in Fig. 1 for Σ_S as in (3) for ρ satisfying (4), then capacity can be upper bounded as

$$\mathcal{C} \leq R^{\text{OUT}} = \begin{cases} \frac{1}{2} \log \left(1 + \frac{P}{1+\bar{\rho}c^2} \right) + \frac{9}{4} & M-1 \geq c^2\bar{\rho} \\ \frac{1}{2M} \log(1+P) & M-1 < c^2\bar{\rho} \leq (M-1)(P+1) \\ + \frac{M-1}{2M} \log(\bar{\rho}c^2) + \frac{3}{2} & \\ \frac{1}{2M} \log(1+P) + 2 & \bar{\rho}c^2 > (M-1)(P+1) \end{cases} \quad (19)$$

for $\bar{\rho} = 1 - \max\{0, \rho\}$ and the exact capacity is to within 2.25 bpcu from the outer bound in (19).

Proof: app: Approximate capacity for the M-CCDP with Gaussian independent states ■

The difficulty in extending the result of Th. V.2 to the case

of any correlation matrix Σ_S lays in the fact that, in this case, decoders have different decoding capabilities and therefore there are a number of ways in which the same set of public bits can be transmitted to each receiver. This can be accomplished by varying the time-sharing ratio for the private codeword for each receiver in the scheme in Fig. 4. This optimization quickly becomes untractable and deriving a matching outer bound is challenging.

VI. CONCLUSION

In this paper we study the capacity of the carbon copying onto dirty paper channel with equivalent states, a variation of the classic dirty paper channel in which the transmitted message is decoded at M receivers, each observing a linear combination of the input, Gaussian noise and one of M possible state sequences. These state sequences are non-causally known at the transmitter and are statistically equivalent, being jointly Gaussian-distributed, with unitary variance and identical pairwise correlation. Although inner and outer bounds to the capacity of this channel are available in the literature, no characterization of capacity was known. We derive the capacity of this model to within 2.25 bits-per-channel-use for any channel and any pairwise correlation among the states. In this model capacity can be approached with a rather simple strategy in which the input is composed of the superposition of two codewords: a bottom, common codeword decoded at all users and a top, private codeword decoded at each receiver for a portion $1/M$ of the time and pre-coded against the channel state experienced at the given receiver. The major contribution of the paper is in the derivation of an outer bound which closely approaches this intuitive inner bound.

Despite of our progress, the capacity of the channel in which the states have any jointly Gaussian distribution remains unknown.

REFERENCES

- [1] S. Gel'fand and M. Pinsker, "Coding for channel with random parameters," *Probl. Contr. Inform. Theory*, vol. 9, no. 1, pp. 19–31, 1980.
- [2] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, 1983.
- [3] A. Khisti, U. Erez, A. Lapidoth, and G. Wornell, "Carbon copying onto dirty paper," *IEEE Trans. Inf. Theory*, vol. 53, no. 5, pp. 1814–1827, May 2007.
- [4] C. Nair, A. E. Gamal, and Y.-K. Chia, "An achievability scheme for the compound channel with state noncausally available at the encoder," *arXiv preprint arXiv:1004.3427*, 2010.
- [5] S. Rini and S. Shamai, "The impact of phase fading on the dirty paper channel," *arXiv preprint arXiv:1401.4236*, 2014.
- [6] Y. Steinberg and S. Shamai, "Achievable rates for the broadcast channel with states known at the transmitter," in *Proceedings. International Symposium on Information Theory, 2005. ISIT 2005.*, 2005.
- [7] K. Marton, "A coding theorem for the discrete memoryless broadcast channel," *IEEE Trans. Inf. Theory*, vol. 25, no. 3, pp. 306–311, May 1979.
- [8] Y. Steinberg, "Coding for the degraded broadcast channel with random parameters, with causal and noncausal side information," *Information Theory, IEEE Transactions on*, vol. 51, no. 8, pp. 2867–2877, 2005.
- [9] S. Rini and S. Shamai, "The impact of phase fading on the dirty paper coding channel," in *Information Theory (ISIT), 2014 IEEE International Symposium on*. IEEE, 2014, pp. 2287–2291.
- [10] —, "The carbon copy onto dirty paper channel with statistically equivalent states," *arXiv preprint*, 2016.

APPENDIX A
PROOF OF LEM. I.1

For the matrix in (3) the leading principal minor can be obtained through the matrix determinant lemma as

$$\det((1-\rho)\mathbf{I}_{m,m} + \rho\mathbf{1}_{1,m}\mathbf{1}_{m,1}) = (1-\rho)^m \left(1 + \frac{m\rho}{1-\rho}\right), \quad (20)$$

which is non-negative for

$$\rho \geq -\frac{1}{m-1}. \quad (21)$$

Consequently, all the leading principal minors of the matrix in (3) are positive when

$$\rho \geq \min_m \left\{ -\frac{1}{m-1} \right\} = -\frac{1}{M-1}. \quad (22)$$

Equation (22) together with the fact that ρ is necessarily bonded below one, we obtain the condition in (4).

APPENDIX B
PROOF OF LEM. I.2.

Given the state sequence vector $\mathbf{S}^N = [S_1^N \dots S_M^N]$, we can represent this sequence as being obtained as $\mathbf{S}_1^N + \mathbf{S}_2^N$ where

$$\begin{aligned} \mathbf{S}_1^N &\sim \text{i.i.d. } \mathcal{N}(0, \rho\Sigma_S) \\ \mathbf{S}_2^N &\sim \text{i.i.d. } \mathcal{N}(0, \bar{\rho}\Sigma_S), \end{aligned}$$

for $\mathbf{S}_1 \perp \mathbf{S}_2$ and $\bar{\rho} = 1 - \rho$.

Consider now the channel in which the set sequences \mathbf{S}_2^N is provided as a side information to the transmitter and all the receivers: the capacity of this channel must necessarily be larger than the capacity of the original channel, since this extra knowledge can be ignored. The m^{th} receiver in the enhanced channel can produce the sequence \tilde{Y}_m as

$$\begin{aligned} \tilde{Y}_m^N &= Y_m^N - cS_{2,m}^N \\ &= X^N + cS_{1,m}^N + Z_m. \end{aligned} \quad (23)$$

The sequence \tilde{Y}_m^N in (23) is statistically equivalent to the channel in (1) for

$$\tilde{c} = c\sqrt{\rho} \leq c, \quad (24)$$

where the state sequence \mathbf{S}' is appropriately scaled so that (2) holds. When considering the equivalent channel output \tilde{Y}_m^N , the sequences in $\tilde{\mathbf{S}}_2$ acts as a common information between the transmitter and the receivers and thus does not increase capacity. From these observations, we conclude that the capacity of the model with state gain c and side information $\tilde{\mathbf{S}}_2$ is equivalent to the capacity of the channel model in which the state gain is \tilde{c} . This implies that the capacity increases as c decreases and thus concludes the proof.

APPENDIX C
PROOF OF TH. III.1.

A gap of 1 bpcu for $P \leq 3$ or $c^2 \leq 3$ can be attained by either treating the state as noise or simple considering the trivial achievable point $R = 0$, so we consider here only the case $P > 3$ and $c^2 > 3$.

The outer bound derivation initially follows steps similar to that of [3, Th. 3] and is successively improved by using of the observation in Lem. I.2. The inner bound is substantially the same inner bound as in [9] and relies on the superposition coding and binning: a base codeword treats the state as noise and two top codewords which are each transmitted only for half of the time. The first codeword is pre-coded against the channel state observed at one user while the second codeword is pre-coded against the state observed at the second user.

Capacity outer bound: As in [3, Th. 3], we have that the capacity of this channel can be upper bounded as

$$N(R - \epsilon) \leq \min_{j \in \{1,2\}} I(Y_j^N; W) \quad (25a)$$

$$\leq \frac{1}{2} (H(Y_1^N) + H(Y_2^N) - H(Y_1^N|W) - H(Y_2^N|W)). \quad (25b)$$

The sum of the positive entropy terms $H(Y_1^N) + H(Y_2^N)$ can be bounded as

$$H(Y_1^N) + H(Y_2^N) \quad (26a)$$

$$\leq \frac{N}{2} \log(P + c^2 + 2c\sqrt{P} + 1) + \frac{N}{2} \log(P + c^2 + 2c\sqrt{P} + 1) + N \log 2\pi e \quad (26b)$$

$$\leq N \log 2\pi e(P + c^2 + 1) + N + N \log 2\pi e, \quad (26c)$$

where (26b) follows from the Gaussian Maximizes Entropy (GME) property and (26c) follows from the fact that

$$2(P + c^2) \geq (\sqrt{P} + c)^2. \quad (27)$$

For the sum of negative entropy terms $-H(Y_1^N|W) - H(Y_2^N|W)$ we have

$$-H(Y_1^N|W) - H(Y_2^N|W) \quad (28a)$$

$$\leq -H(Y_1^N, Y_2^N|W) \quad (28b)$$

$$= -H(Y_2^N - Y_1^N, Y_2^N|W) \quad (28c)$$

$$= -H(c(S_1^N - S_2^N) + Z_2^N - Z_1^N, Y_2^N|W), \quad (28d)$$

where in (28c) we have used the transformation

$$\begin{bmatrix} Y_2^N - Y_1^N \\ Y_2^N \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} Y_1^N \\ Y_2^N \end{bmatrix} \quad (29)$$

which has jacobian equal one. We now continue the series of inequalities in (28) as

$$= -H(c(S_2^N - S_1^N) + Z_2^N - Z_1^N|W) - H(Y_2^N|S_2^N - S_1^N + Z_1^N - Z_2^N, W) \quad (30a)$$

$$\leq -H(c(S_2^N - S_1^N) + Z_2^N - Z_1^N) - H(Y_2^N|S_1^N, S_2^N, W, Z_1^N - Z_2^N) \quad (30b)$$

$$\leq -H(c(S_2^N - S_1^N) + Z_2^N - Z_1^N) - H(Z_2^N|Z_1^N - Z_2^N). \quad (30c)$$

Since $Z_1^N \perp Z_2^N$, we obtain

$$\begin{aligned} -H(Y_1^N|W) - H(Y_2^N|W) &\leq N \left(-\frac{1}{2} \log 2\pi e(2c^2 + 2) - \frac{1}{2} \log 2\pi e \frac{1}{2} \right) \\ &= N \left(-\frac{1}{2} \log 2\pi e(c^2 + 1) - \frac{1}{2} \log 2\pi e \right). \end{aligned}$$

The two above inequalities establish the outer bound

$$\begin{aligned} R^{\text{OUT}} &= \frac{1}{2} \log(P + c^2 + 1) \\ &\quad - \frac{1}{4} \log(c^2 + 1) + \frac{1}{2}. \end{aligned} \quad (31)$$

Since the capacity of the channel is decreasing in c^2 , as shown in Lem. I.2, we can optimize the outer bound in (32) over the set $c' \in [0, c]$. In order to match the boundaries of the optimization in the inner and the outer bound, we choose to further loosen the outer bound in (31) to

$$\begin{aligned} R^{\text{OUT}} &= \frac{1}{2} \log(P + c^2 + 1) \\ &\quad - \frac{1}{4} \log(c^2) + \frac{1}{2}. \end{aligned} \quad (32)$$

The first derivative of (32) in c^2 is

$$\frac{\partial (32)}{\partial c^2} = -\frac{1}{4} \frac{P + 1 - c^2}{(1 + P + c^2)c^2}, \quad (33)$$

which has a zero in $c^* = \sqrt{P + 1}$. For $c^2 = P + 1$, the second derivation of (32) in c^2 is positive: we can therefore set $c' = \min\{\sqrt{P + 1}, c\}$ and obtain a channel with a larger capacity but a tighter expression of the outer bound in (32). The result of the optimization in c correspond to bound in (10). Note that, for the case $c^2 < 1$ we use the trivial outer bound $C \leq \frac{1}{2} \log(P + 1)$: since the variance of the state is 1, the contribution of the state to the channel output is minimal for this case.

Capacity inner bound: Consider the transmission scheme in which the channel input X^N is comprised of the superposition of the following codewords:

- (i) the base codeword X_{SAN}^N (SAN as in “State As Noise”) which treats the state as noise and
- (ii) the top codewords $X_{\text{PAS}-i}^N$ (PAS as in “Pre-coded Against the State”) is pre-coded against the sequence S_i^N for $i \in \{1, 2\}$. Additionally $X_{\text{PAS}-1}^N$ is transmitted for the first half of the time, while $X_{\text{PAS}-2}^N$ is transmitted for the second

half of the time. The codewords $X_{\text{PAS}-i}^N$ are superimposed over the codeword X_{SAN}^N : receiver i jointly decodes X_{SAN}^N and $X_{\text{PAS}-i}^N$. All the codewords are iid Gaussian-distributed: X_{SAN}^N has power αP while $X_{\text{PAS}-i}$ have both power $\bar{\alpha}P$ for some $\alpha \in [0, 1]$, $\bar{\alpha} = 1 - \alpha$. The common codeword X^{SAN} attain the rate

$$\begin{aligned} R_{\text{SAN}} &= I(X_{\text{SAN}}; Y_i) \\ &= \frac{1}{2} \log \left(1 + \frac{\alpha P}{1 + c^2 + \bar{\alpha}P} \right) \end{aligned} \quad (34)$$

and is used to communicate the messages $W_{\text{SAN}} \in [1 \dots 2^{NR_{\text{SAN}}}]$ to both users simultaneously. The two private codewords $X_{\text{PAS}-1}$ and $X_{\text{PAS}-2}$ each attain the rate

$$\begin{aligned} R_{\text{PAS}} &= I(Y_1; U_1 | X_{\text{SAN}}) - I(U_1; S_1) \\ &= I(Y_2; U_2 | X_{\text{SAN}}) - I(U_2; S_2) \end{aligned} \quad (35)$$

and encode the same message $W_{\text{PAS}} \in [1 \dots 2^{NR_{\text{PAS}}}]$. Note that the message W_{PAS} is sent twice, since it is reliably communicated to the first users in the first half of the transmission and to the second user the second half of the transmission. Combining the rate of the common and the private message, we conclude that the overall attainable rate is

$$R^{\text{IN}}(\alpha) = \frac{1}{2} \log \left(1 + \frac{\alpha P}{1 + \bar{\alpha}P + c^2} \right) + \frac{1}{4} \log(1 + \bar{\alpha}P), \quad (36)$$

for any $\alpha \in [0, 1]$. The optimization over α yields that the optimal value

$$\bar{\alpha}^* = \begin{cases} 0 & c^2 < 1 \\ \frac{c^2 - 1}{P} & 1 \leq c^2 < P + 1 \\ 1 & c^2 \geq P + 1 \end{cases} \quad (37)$$

and the corresponding optimal rates

$$R^{\text{IN}} = \begin{cases} \frac{1}{2} \log \left(1 + \frac{P}{c^2 + 1} \right) & c^2 < 1 \\ \frac{1}{2} \log(1 + c^2 + P) - \frac{1}{4} \log(c^2) - \frac{1}{2} & 1 \leq c^2 < P + 1 \\ \frac{1}{4} \log(P + 1) & c^2 \geq P + 1 \end{cases} \quad (38)$$

Gap between inner and outer bound: For the case $c^2 \leq 1$ we notice that the distance between inner and outer bound is at $1/2$ bpcu using simple considerations on the shape of the capacity region. For the remaining cases, inner and outer bounds can be compared directly: the gap is 1 bpcu for $c^2 \geq P + 1$ and also 1 bpcu for the case $1 \leq c^2 < P + 1$. We therefore conclude that, regardless of the channel parameters, the outer bound can be attained to within 1 bpcu.

APPENDIX D PROOF OF TH. IV.1.

The proof is an extension of the proof of Th. III.1 and thus relies on similar inner and outer bounding techniques. The first part of the outer bound derivation follows the derivation of [3, Eq. (31)] but later we employ a recursive bounding of the mutual information terms to come to a tighter bounding. On the other hand the inner bound is a rather straight forward extension of the bound in Th. III.1 in which a bottom codeword and multiple top, pre-coded codewords are used to communicate the common message.

As for the proof in App. C, we only need to consider the case $P > 3$ and $c^2 > 3$ since the capacity region is smaller than 1bpcu otherwise.

Capacity outer bound: As in [3, App. 3.C], we write

$$N(R - \epsilon) \leq \min_{m \in [1 \dots M]} I(Y_m^N; W) \quad (39a)$$

$$\leq \frac{1}{M} \sum_{m=1}^M I(Y_m^N; W) \quad (39b)$$

$$\leq \max_{m \in [1 \dots M]} H(Y_m^N) - \frac{1}{M} \sum_{i=1}^M H(Y^N | W) \quad (39c)$$

$$\leq \frac{N}{2} \log(P + c^2 + 2c\sqrt{P} + 1) + \frac{N}{2} \log(2\pi e) - \frac{1}{M} \sum_{m=1}^M H(Y_m^N | W). \quad (39d)$$

We now proceed in establishing a recursion by defining the term T_m as

$$T_m \triangleq \sum_{i=m}^M H(Y_m^N | W), \quad (40)$$

which allows us to rewrite (39d) as

$$N(R - \epsilon) \leq \frac{N}{2} \log(P + c^2 + 2c\sqrt{P} + 1) + \frac{N}{2} \log(2\pi e) - T_1. \quad (41)$$

The term T_1 can now be rewritten as

$$-T_1 = -H(Y_1^N | W) - H(Y_2^N | W) - T_3. \quad (42)$$

We have seen in (28) that the difference $-H(Y_1^N | W) - H(Y_2^N | W)$ can be bounded as follows:

$$-T_1 \leq -H(c(S_1^N - S_2^N) + Z_2^N - Z_1^N, Y_2^N | W) - T_3. \quad (43)$$

Since the noise terms are Z_i to be i.i.d. and identically distributed we have:

$$-T_1 = -H(c(S_1^N - S_2^N)) - H(Y_2^N | S_1^N - S_2^N, W) - H(Y_3^N | W) - T_4 \quad (44a)$$

$$= -\frac{N}{2} \log(2c^2) - H(Y_2^N | S_1^N - S_2^N, W) - H(Y_3^N | W) - T_4 \quad (44b)$$

$$\leq -\frac{N}{2} \log(2c^2) - H(Y_2^N, Y_3^N | S_1^N - S_2^N, W) - T_4 \quad (44c)$$

$$\leq -\frac{N}{2} \log(2c^2) - H(Y_3^N - Y_2^N, Y_3^N | S_1^N - S_2^N, W) - T_4 \quad (44d)$$

$$\leq -\frac{N}{2} \log(2c^2) - H(c(S_3^N - S_2^N), Y_3^N | S_1^N - S_2^N, W) - T_4 \quad (44e)$$

$$= -\frac{N}{2} \log(2c^2) - H(c(S_3^N - S_2^N) | S_1^N - S_2^N) - H(Y_3^N | S_1^N - S_2^N, S_3^N - S_2^N, W) - T_4 \quad (44f)$$

$$\leq -\frac{N}{2} \log(2c^2) - \frac{N}{2} \log\left(\frac{3}{2}c^2\right) - H(Y_3^N | S_1^N - S_2^N, S_3^N - S_2^N, W) - T_4. \quad (44g)$$

A recursion can now be established on the same lines as (44) to bound all the terms in the summation T_1 : let $\Delta_1^N = \mathbf{0}_{1,N}$ and define Δ_i^N , $i > 1$ as

$$\Delta_i^N \triangleq S_i^N - S_{i-1}^N, \quad (45)$$

then we can write

$$T_1 \leq \sum_{i=2}^K H(c\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N) - H(Y_K^N | \Delta_1^N \dots \Delta_K^N) - T_{K+1} \quad (46a)$$

$$= \sum_{i=2}^K H(c\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N) - H(Y_K | \Delta_1^N \dots \Delta_K^N, W) - H(Y_{K+1} | W) - T_{K+2} \quad (46b)$$

$$\leq \sum_{i=2}^K H(c\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N) - H(Y_K, Y_{K+1} | \Delta_1^N \dots \Delta_K^N, W) - T_{K+2}. \quad (46c)$$

By proceeding in this manner up to $K = M$ we come to the bound

$$-T_1 \leq \sum_{i=2}^M -H(c\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N) - H(Y_M^N | \Delta_1^N \dots \Delta_M^N, W) \quad (47a)$$

$$\leq \sum_{i=2}^M -H(c\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N) - H(Z_M^N) \quad (47b)$$

$$\leq \sum_{i=2}^M -H(c\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N) - \frac{N}{2} \log(2\pi e). \quad (47c)$$

We are now left to evaluate the intermediate terms in the summation:

$$H(c\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N) = \frac{1}{2} \log(c^2) + H(\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N). \quad (48)$$

The correlation matrix among the entries of the vector $[\Delta_2^N \dots \Delta_M^N]$ is

$$\Sigma_{\Delta^N} = \begin{bmatrix} 2 & -1 & 0 & 0 & \dots & 0 \\ -1 & 2 & -1 & & \dots & 0 \\ 0 & -1 & 2 & -1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ & & 0 & -1 & 2 & -1 & 0 \\ \vdots & & & 0 & -1 & 2 & -1 \\ 0 & \dots & & & 0 & -1 & 2 \end{bmatrix} \quad (49)$$

and thus we conclude that

$$-H(\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N) = -\frac{N}{2} \log \left(2 - [-1 \dots -1] \cdot \begin{bmatrix} 2 & -1 & 0 \\ -1 & \ddots & \ddots \\ 0 & \ddots & \end{bmatrix} \cdot \begin{bmatrix} -1 \\ \vdots \\ -1 \end{bmatrix} \right) \quad (50a)$$

$$= -\frac{N}{2} \log \left(2 - \frac{i-1}{i} \right) \quad (50b)$$

$$\leq -\frac{N}{2} \log 1 = 0, \quad (50c)$$

where (50b) follows from the properties of symmetric tri-diagonal matrices.

With the bounding in (50), we obtain the outer bound

$$\begin{aligned} R^{\text{OUT}} &\leq \frac{1}{2} \log(1 + P + c^2) - \frac{M-1}{2M} \log c^2 + \frac{1}{2} \log 2\pi e \left(1 - \frac{1}{2M} \right) \\ &\leq \frac{1}{2} \log(1 + P + c^2) - \frac{M-1}{2M} \log c^2 + \frac{3}{2}, \end{aligned} \quad (51)$$

and, as for the proof of Th. III.1, this outer bound can be optimized over c in the interval $c' \in [0, c]$. The derivative of (51) in c^2 is equal to zero in

$$c^* = \sqrt{(M-1)(P+1)}, \quad (52)$$

while the second derivative is positive in this point. Having that $c^{2*} = (M-1)(P+1)$ is a minimum of the outer bound in (51) when $c^2 \geq (M-1)(P+1)$, we obtain the outer bound expression in (51) for $c^2 > M-1$.

For the interval $M-1 \geq c^2$ we bound the expression in (51) as follows:

$$\begin{aligned} &\frac{1}{2} \log(1 + P + c^2) - \frac{M-1}{2M} \log c^2 + \frac{3}{2} \\ &\leq \frac{1}{2} \log \left(1 + \frac{P}{1+c^2} \right) + \frac{1}{2} \log(1 + c^2) - \frac{M-1}{2M} \log c^2 + \frac{3}{2} \end{aligned} \quad (53a)$$

$$\leq \frac{1}{2} \log \left(1 + \frac{P}{1+c^2} \right) + \frac{1}{2} \log(2c^2) - \frac{M-1}{2M} \log c^2 + \frac{3}{2} \quad (53b)$$

$$= \frac{1}{2} \log \left(1 + \frac{P}{1+c^2} \right) + \frac{1}{2M} \log(c^2) + 2 \quad (53c)$$

$$= \frac{1}{2} \log \left(1 + \frac{P}{1+c^2} \right) + \frac{1}{2M} \log(M-1) + 2 \quad (53d)$$

$$\leq \frac{1}{2} \log \left(1 + \frac{P}{1+c^2} \right) + \frac{1}{4} + 2, \quad (53e)$$

where (53b) follows from the assumption that $c^2 > 1$ and (53e) from the fact that $x^{-1} \log(x-1)$ has a maximum in $x = 4$ when x is integer valued. Combining these results, we obtain the desired outer bound in (10).

Capacity inner bound: Consider an inner bound which extends the inner Th. III.1 and in which the inner bound is composed of the superposition of two codewords:

- (i) the base codeword X_{SAN}^N (*SAN* as in “*State As Noise*”) which treats the state as noise and
- (ii) the top codewords $X_{\text{PAS-i}}^N$ (*PAS* as in “*Pre-coded Against the State*”) is pre-coded against the sequence S_i^N for

$i \in \{1 \dots M\}$, each transmitted only for a portion $1/M$ of the time. The rate achieved by each user with this scheme is

$$R^{\text{IN}} = \max_{\alpha \in [0,1]} \frac{1}{2} \log \left(1 + \frac{\alpha P}{1 + \bar{\alpha} P + c^2} \right) + \frac{1}{2M} \log(1 + \bar{\alpha} P). \quad (54)$$

The optimization over α yields the achievable rate

$$R^{\text{IN}} = \begin{cases} \frac{1}{2} \log \left(1 + \frac{P}{1+c^2} \right) & M-1 > c^2 \\ \frac{1}{2} \log(P + c^2 + 1) & M-1 \leq c^2 \leq (M-1)(P+1) \\ -\frac{M-1}{2M} \log(c^2) - \frac{1}{2} & \\ \frac{1}{2M} \log(1+P) & \end{cases} \quad (55)$$

Gap between inner and outer bound: Consider the case where $M > 2$, and compare the expression in (51) and in (55). The gap for $M-1 > c^2$ is at most 2 bpcu, for the case $M-1 \leq c^2 \leq (M-1)(P+1)$ and it is 2 bpcu also when $c^2 > (M-1)(P+1)$ is $1/2M$. The largest gap between inner and outer bound is 2.25 bpcu and is attained for $M-1 \leq c^2$.

APPENDIX E PROOF OF TH. V.1

Lets consider the case of positive and negative correlation separately, since they require a separate derivation

Approximate capacity for $\rho > 0$: For the outer bound we simply consider the outer bound in (10) obtained by providing S_c to both decoders: after this term is stripped from the channel output, the receivers obtain the same model as in Th. III.1 but with a state with smaller variance, that is $1 - \rho$ instead of 1. By absorbing this factor in c , we obtain the outer bound in (19).

For the inner bound consider the generalization of the inner bound in App. C in which the base codeword X_{SAN}^N is pre-coded against the state S_c^N so that the rate is can be transmitted at rate

$$R^{\text{SAN}} = \frac{1}{2} \log \left(1 + \frac{\alpha P}{1 + \bar{\alpha} P + (1 - \rho)Q} \right). \quad (56)$$

With adjustment to the attainable scheme in Th. III.1, we see that the region in (16) can be attained to within 1 bpcu.

Approximate capacity for $\rho < 0$.

Note that the correlation affects the derivation of the outer bound in App. C only in the derivation of the term (30c) where in can be noted that $H(c(S_2^N - S_1^N) + Z_2^N - Z_1^N)$ is decreasing in the correlation ρ

$$H(c(S_2^N - S_1^N) + Z_2^N - Z_1^N) = \frac{1}{2} \log(2c^2(1 - \rho)), \quad (57)$$

accordingly we have that outer bound for independent states in an outer bound for the case of negatively correlated states. The negative correlation, also, does not affect the inner bound in Th. III.1 so that the same rate as in (38) is attainable. With these two considerations we see that the capacity for the case of negative correlation can be approached in the same manner as the case of independent states.

APPENDIX F PROOF OF TH. V.2

The case of positive correlation straightforwardly extends from the proof of Th. V.1 in App. E.

For the case of negative correlation, we shall show that the recursion in (48) is not affected by the negative correlation and that the value of the entropy term in (50) are decreasing in the value of the correlation.

Note that the covariance matrix in (49) is not affected by the correlation, since

$$\begin{aligned} \text{Var}[\Delta_i] &= \text{Var}[S_i - S_{i-1}] \\ &= 2(1 - \rho), \end{aligned}$$

and equivalently

$$\begin{aligned} \mathbb{E}[\Delta_i \Delta_{i+1}] &= \mathbb{E}[(S_i - S_{i-1})(S_{i+1} - S_i)] \\ &= \rho - 1 - \rho + \rho \\ &= -(1 - \rho), \end{aligned}$$

while for $j > i + 1$

$$\mathbb{E}[\Delta_i \Delta_j] = \mathbb{E}[(S_i - S_{i-1})(S_j - S_{j-1})]$$

$$\begin{aligned}
&= \rho - \rho - \rho + \rho \\
&= 0,
\end{aligned}$$

so that

$$\begin{aligned}
-H(\Delta_i^N | \Delta_1^N \dots \Delta_{i-1}^N) &= -\frac{N}{2} \log \left((1 - \rho)^N \left(2 - \frac{i-1}{i} \right) \right) \\
&\leq -\frac{N}{2} \log 1 = 0,
\end{aligned} \tag{58}$$

where the expression in (58) is increasing in ρ and thus once again obtain that the outer bound for negative correlation is upper bounded by the outer bound for independent states. As in the proof of Th. V.1 in App. E, the inner bound is not affected by negative correlation: we therefore conclude that the capacity for the case of negative correlated states is bounded in the same manner as in Th. V.2.